International Journal of Advanced Smart Convergence Vol.8 No.1 133-140 (2019)
http://dx.doi.org/10.7236/IJASC.2019.8.1.133

IJASC 19-1-16

An Evidence Retraction Scheme on Evidence Dependency Network

Gye Sung Lee

Dept. of Software, Dankook University, Korea
gslee@dankook.ac.kr

Abstract

In this paper, we present an algorithm for adjusting degree of belief for consistency on the evidence dependency network where various sets of evidence support different sets of hypotheses. It is common for experts to assign higher degree of belief to a hypothesis when there is more evidence over the hypothesis. Human expert without knowledge of uncertainty handling may not be able to cope with how evidence is combined to produce the anticipated belief value. Belief in a hypothesis changes as a series of evidence is known to be true. In non-monotonic reasoning environments, the belief retraction method is needed to clearly deal with uncertain situations. We create evidence dependency network from rules and apply the evidence retraction algorithm to refine belief values on the hypothesis set. We also introduce negative belief values to reflect the reverse effect of evidence combination.

Keywords: *Dempster-Shafer theory, Evidence Combination, Non-monotonic Reasoning, Belief Function, Focal Element, Basic Probability Assignment, Mass Function.*

1. Introduction

Uncertainty handling, a traditional AI topic, plays an important role in reasoning and knowledge based systems. There are many different approaches to handling uncertainty [1,2]. In the probabilistic approach, formal probability theory is the basis of determining the possibility of an event or an outcome. Bayesian inference deals with uncertainty and requires updating the probability for a hypothesis as more evidence or information becomes available [3,4]. Fuzzy logic analysis extends traditional two-valued logic to be a continuous logic valued from 0 to 1 [5,6]. The Dempster-Shafer theory, also referred to as the theory of belief functions or evidence theory, is added up to the field of uncertainty handling [2,8]. Uncertainty comes from situations where information is limited or unavailable due to lack of knowledge or insufficient information. More evidence provided into the system is combined so that the current set of hypotheses with belief values will be updated by the rule of combination. Among other uncertainty handling schemes, the evidence combination of the Dempster-Shafer theory provides unique features that differentiates it from others.

Manuscript Received: Feb. 7, 2019 / Revised: Feb. 12, 2019 / Accepted: Feb. 16, 2019
Corresponding Author: gslee@dankook.ac.kr
Tel: +82-10-3344-7529, Fax: +82-31-8005-3223
Dept. of Software, Dankook University, Korea

The Dempster-Shafer theory, like other schemes, has been used for reasoning with degree of belief to measure how evidence is supportive of hypothesis [2,8]. The theory involves the assigning of degree of belief over sets of hypotheses and computing the combined effect on the hypotheses sets using Dempster's rule of combination. In reasoning with uncertainty by the Dempster-Shafer scheme, the belief in a given hypothesis is constantly changing as a sequence of evidence comes into play in the process of reasoning [8-11]. The result of evidence combination may lead to confirmation of a hypothesis, but also dissolve the inconsistencies and/or conflicts during the process of evidence combination. For example, there is a hypothesis set, H = {meningitis, brain tumor} in which the belief value for meningitis to be 0.5. The rest of the belief value, 0.5, is assigned to the whole set, H. If another piece of evidence becomes available and supports brain tumor, with a belief value of 0.5, the new evidence is combined with the existing one resulting in a belief value of 1/3 each for meningitis and brain tumor. They are both likely with a reduced degree of belief. In general, most of uncertainty handling schemes are based on the supportive reasoning approach, in which evidence should support the hypothesis. However, we can imagine a situation where another piece of evidence can refute the same hypothesis. It could be the case, where by some reason, if the evidence supporting brain tumor should not have been asserted. The support for brain tumor should be retracted from the belief value in brain tumor, and belief values for meningitis and brain tumor should be recalculated. These types of problems can easily happen in real world applications and is well studied in the field of non-monotonic reasoning with propositional or predicate logics [7].

In building knowledge-based systems, it is improbable to expect that all the knowledge needed for reasoning tasks could be acquired at the outset. It is even more unlikely that the knowledge acquired is to be complete. Traditionally, we start with an initial set of knowledge that is typically incomplete and contains a lot of redundancies, inconsistencies, and other types of disparities. Even if we can build up complete and valid knowledge initially, it would not be possible to keep it valid in a continually changing environment.

During the process of building knowledge for an application through the domain expert, there has always been a gap between the system engineer and human expert due to the difference in perspectives over the way of reasoning. Consider a case where two rules are related by subsumption, where one rule is subsumed by another. It occurs when all the evidence (condition part) of a rule is completely contained in the evidence of the other rule. These types of rules are not recommended in rule-based systems. But to human experts, it is common in their way of reasoning. When more specific evidence is available, higher degree of belief is assigned to the hypothesis. Here the difference between the system engineer and human expert sometimes poses a danger in determining the degree of belief by evidence combination. Let's assume that there are two rules in subsumption relations, and a more specific set of conditions are met. Two rules are ready to fire and determination by evidence combination could lead to higher belief value than the human expert estimates. As a simple example, the first rule includes one condition, a_1, and concludes x with a belief value of 0.4. The second rule comes with two conditions, a_1, a_2, and concludes x with a belief value of 0.7. The first rule is subsumed by the second rule. By the second rule, the author of the rules implies higher belief value and assigns a degree of belief of 0.7 to the hypothesis, x. In the evidence combination scheme, when two conditions, $a_1\ and\ a_2$, are known to be true, the conclusion results in x with a new belief value of 0.82. It would be a little higher than it should be if the original intention of the expert is to assign a belief value of 0.7. The belief value of the rule should be adjusted to make the resultant value be 0.7. As the size of the knowledge base grows, the relationships among evidence become more complicated. A complex evidence dependency network is constructed and a method to refine belief values for proper evidence combination should be followed. The rest of the paper is organized as follows. In section 2, basic principles of Dempster-Shafer theory are discussed. In section 3, the evidence retraction scheme that reverses the effect of evidence combination is described. In section 4, several cases are introduced and belief values for different cases are refined with the method. Section 5 concludes the paper.

2. Dempster-Shafer Evidence Combination Scheme

The Dempster-Shafer (DS) theory is a mathematical theory of evidence, first introduced by Arthur P. Dempster and developed by Glenn Shafer[2,8]. The DS theory can be interpreted as a generalization of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons. Evidence can be associated with multiple possible events or hypotheses. The theory includes the Transferable Belief Model which obtains degrees of belief for one question from subjective probabilities for a related question. This includes the rule of combination which governs the combining probabilities when independent pieces of evidence become available. Since the DS theory does not require an assumption regarding the probability of the individual constituents of the set or interval, the DS theory is a potentially valuable tool for the evaluation of risk and reliability in engineering applications, where it is not possible to obtain a precise measurement from experiments, or when knowledge is obtained from expert elicitation. The Dempster-Shafer theory begins with a frame of discernment (Θ) which is a finite set of all possible mutually-exclusive hypotheses. Subjective probabilities (masses) can be assigned to all subsets of the frame. This is called basic probability assignment function (bpa), represented by m.

$$\mathrm{m}: 2^{\Theta} \to [0,1] \tag{1}$$

In Eq. (1), 2^{Θ} is the power set of the frame of discernment, Θ. The function, m, is called a mass function with two axioms: the mass of empty set (ϕ) is zero, and the sum of masses for all members of 2^{Θ} is 1. Assume that A is a set in the power set of Θ.

$$\mathrm{m}(\phi) = 0,\ \sum_{A \subseteq \Theta} m(A) = 1 \tag{2}$$

In Eq. (2), the measure, $\mathrm{m}(\mathrm{A})$, quantifies the proportion of all relevant and available evidence that supports A. A set that has non-zero mass is called a focal element. From the basic probability assignment, the upper and lower bounds can be defined, plausibility and belief measure, respectively. The lower bound Belief for set A is defined in Eq. (3) as the sum of all the basic probability assignments of the proper subsets of the set A, i.e., for $\mathrm{B} \subseteq \mathrm{A}$.

$$\mathrm{Bel}(\mathrm{A}) = \sum_{B \subseteq A} m(B) \tag{3}$$

Belief in a hypothesis expresses the amount of belief that directly supports either the given hypothesis or a more specific one. It ranges from 0 to 1. The plausibility measure is defined in Eq. (4) as the sum of all the basic probability assignments off the set that intersect the set of A.

$$\mathrm{Pl}(\mathrm{A}) = \sum_{B \cap A \neq \phi} m(B) = 1 - Bel(\bar{A}) \tag{4}$$

$\bar{A}$ is set complement of A.

Evidence combination needs to combine two independents set of basic probability assignments. The joint mass is calculated from the two sets of masses, m_1 and m_2 as Eq. (5).

$$m_{1,2}(A) = (m_1 \oplus m_2)(A) = \frac{1}{1-K} \sum_{B \cap C = A \neq \phi} m_1(B) m_2(C) \tag{5}$$

where K, representing basic probability mass associated with conflict, is defined as follows:

$$\mathrm{K} = \sum_{B \cap C = \phi} m_1(B) m_2(C) \tag{6}$$

The denominator in the joint mass, $1 - \mathrm{K}$, is used as a normalization factor by which the conflict is completely ignored.

3. Retracting Dempster-Shafer Evidence Eombination

The Dempster-Shafer scheme includes the rule of evidence combination that combines two belief functions to generate a new belief function, i.e.,

$$Bel_p \oplus Bel_q = Bel_r \quad (7)$$

where Bel_p is assumed to be a belief function for evidence under consideration and Bel_q is the belief function for the rest of evidence that was already combined so far. The focal elements (FEs) of Bel_p is represented by $\{p_i | 1 \leq i \leq l\}$ and mass function, m_p. The focal elements of Bel_q, FE_q is represented by $\{q_i | 1 \leq i \leq m\}$ and mass function m_q. The focal elements of Bel_r is represented by $\{r_i | 1 \leq i \leq n\}$ and mass function m_r.

$$FE_p = \{p_i, a_i | m_p(p_i) = a_i,\ 1 \leq i \leq l\} \quad (8)$$

$$FE_q = \{q_i, b_i | m_q(q_i) = b_i,\ 1 \leq i \leq m\} \quad (9)$$

$$FE_r = \{r_i,\ c_i | m_r(r_i) = c_i,\ 1 \leq i \leq n\} \quad (10)$$

The $\oplus$ operator is used to represent the evidence combination between two belief functions. When evidence for Bel_p is retracted from an evidence pool, its effect contributed to the degree of belief in the current hypothesis set should also be retracted by reversing the evidence combination process. Retracting Bel_p from Bel_r as shown in Eg. (11) recovers the belief state back to the state of Bel_q.

$$Bel_q = Bel_r \ominus Bel_p \quad (11)$$

where $\ominus$ represents an evidence retraction operator.

The Expert's rule is composed of two parts: E, evidence set and H, hypothesis set.

Rule: E → H with degree of belief for h_i

$\mathrm{E} = \{e_i |\ evidence\ related\ to\ H, 1 \leq i \leq s\}$

$\mathrm{H} = \{h_i |\ hypothesis\ concluded\ from\ E, 1 \leq i \leq t\}$

From E, the relations among evidence sets construct an evidence network. For example, assume that there are four sets of evidence, $\{\{e_1, e_2\}, \{e_1, e_4\}, \{e_1, e_2, e_3\}, \{e_1, e_2, e_4\}\}$. When two sets are connected by subset relation, links with direction from subset to superset are drawn. The evidence network for four sets of evidence is in Figure 1. The hypothesis network would be constructed in the same way.

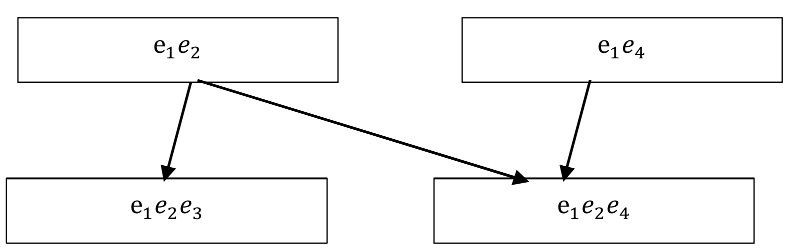

Figure 1. Evidence Dependence Network

After evidence combination, the resultant Bel_r includes the focal element set of Bel_p and Bel_q. It also includes $\{p_1q_1, p_1q_2, ...\}$, where p_iq_j is a non-empty set with non-zero belief value resulting from set intersection of p_i and q_j. This implies that Bel_r always includes the non-empty intersection set of p_i and q_j as well as p_i's and q_j's. Therefore, Bel_p and Bel_q are subsets of Bel_r. As new evidence is asserted, evidence combination increases the size of the focal element set, $FE_p \subseteq FE_q \subseteq FE_r$. Table 1 shows the Dempster's combination. $a_i's$ and $b_i's$ are belief values for $p_i's$ and $q_i's$, respectively and $c_{ij} = a_i \cdot b_j$. The p_iq_j term represents set intersection of $p_i \in FE_p$ and $q_j \in FE_q$. The focal elements that exist in FE_r but not in FE_q are added with zero belief value to Bel_q. This makes the reverse evidence combination computation possible because Bel_q is a subset of Bel_r and the amount of effect for that particular set of focal elements should be identified. The number of focal elements of Bel_q should be n.

Table 1. Dempster's Combination

	Bel_q			
	$q_1(b_1)$	$q_2(b_2)$	...	$q_n(b_n)$
$p_1(a_1)$	$p_1q_1(c_{11})$	$p_1q_2(c_{12})$	...	$p_1q_n(c_{1n})$
$p_2(a_2)$	$p_2q_1(c_{21})$	$p_2q_2(c_{22})$	...	$p_2q_n(c_{2n})$
Bel_p ...			...	
$p_l(a_l)$	$p_lq_1(c_{l1})$	$p_lq_2(c_{l2})$	...	$p_lq_n(c_{ln})$

Note that $q_i = r_i$ for $1 \leq i \leq n$. p_iq_j could be a null set or a member of focal element of Bel_r. Belief value, c_k, for a focal element, r_k, of Bel_r can be computed as follows:

$$Bel_r(c_k) = \frac{1}{1-K}\sum_{i=1}^{l}\sum_{j=1}^{n} c_{ij} \text{ for } p_iq_j = c_k \tag{12}$$

where

$$\mathrm{K} = \sum_{i=1}^{l}\sum_{j=1}^{n} c_{ij} \; for \; p_iq_j = \phi \tag{13}$$

To find the belief values of Bel_q, n equations for Bel_r are rearranged for p_i. These form a set of linear equations. Solving this set of equations produces the belief values for the focal elements of Bel_q. The given p_i's and r_i's, simultaneous equations are represented by $\mathbf{C} \cdot \mathbf{q} = \mathbf{r}$, where $\mathbf{C}$ is a coefficient matrix obtained from these equations and $\mathbf{q}$ is a variable vector composed of q_i's. r is a constant vector given by r_i's. As a result of set intersection, if it is empty, it is summed up separately. If c_{ij} belongs to c_k, $a_j \cdot c_i$ is summed up to a coefficient for c_k. Figure 2 shows how coefficient matrix, C, is computed.

```
for k=1, n {              // for c_k
    for i=1, n {          // for b_i
        coeff_ci = 0.0;
        coeff_null = 0.0;
        for j=1, l   {    // for a_j
            if d_ij == φ
                coeff_null += p_j · r_i;
            else if d_ij == c_i
                coeff_ci += p_j;
        }
        c_ki=coeff_ci + coeff_null;
    }
}
```

Figure 2. Building Coefficient Matrix

4. Belief Value Adjustment

It is very likely in the real world that as more evidence become available, more specific conclusions can be drawn. More supporting evidence sometimes increases degree of belief in a hypothesis. Extra evidence may cut off some hypotheses under consideration. When various evidence is related, an evidence network is formed. In this case, belief adjustment would be a more difficult task. The evidence retraction scheme is applied to three types of situations:

1) increase in belief value when more specific evidence becomes available.

2) exclusion of a hypothesis when more specific evidences become available.

3) belief value adjustment over an evidence network.

A simple example is shown below:

R_1: if a then x with 0.4

R_2: if a and b then x with 0.7

When evidence a is the only evidence available at the time, the subjective belief in x remains low with a degree of belief, 0.4, as designed by R_1. When extra evidence, b, becomes available, R_2 will fire and belief on x will be 0.82 by Dempster's rule of combination. However, human expert suggests that when a and b are confirmed, the degree of belief would be 0.7. Since R_2 is subsumed by R_1, when two independent evidences are true, human expert expects that only R_2 should fire. In this case, this problem should be resolved. One possible solution is somehow to block R_1 from firing when a and b are available. This solution is not a good idea simply because it could not draw a conclusion until all the evidence is fed into the system. Furthermore, it needs a more complex control structure for reasoning.

The second option for solving this problem is to retract the rule firing by the evidence retraction method discussed in section 3. The portion of R_1 could be retracted from a line of reasoning. This solves the problem but requires additional tests for subsumption during the reasoning process. It also wastes processing time in order to undo the evidence combination. The third option is to provide an adjusted belief value for R_2 when the rule is constructed. Evidence combination will produce the exact amount of belief value, 0.7, when both evidences are asserted and two rules are fired. R_2's belief value is changed to 0.5:

R_2: if a and b then x with 0.5

When evidence a and b are asserted, R_1 and R_2 are fired and the belief value will be 0.7 by regular evidence combination as the human expert suggests. This adjustment should be done just once when the rule is initially introduced to the system by building an evidence network and checking the subsumption test over neighboring nodes in the network. Since the discovery of subsumed rules is beyond the scope of this paper, no further details will be discussed.

As another case for the evidence retraction algorithm, a rule includes multiple hypotheses. A set of multiple hypotheses on focus are narrowed down to a smaller set of hypotheses as more evidence becomes available. In this case, some of the hypotheses should be reduced to a belief value of zero so that the hypotheses themselves are removed from the rule. For example, as shown below, R_3 states that if a and b are confirmed, two hypotheses are believed to be true, with a degree of belief in x, 0.5 and a degree of belief in y, 0.2. R_4 states that when extra evidence, c, is available, the belief value for x increases to 0.6 and y should no longer be true and removed from the conclusion.

R_3: if a and b then x with 0.5, y with 0.2

R_4: if a, b and c then x with 0.6

When three conditions are met, both rules fire. The results are x with 0.7727 and y with 0.0909. As discussed above, the human expert's intention is that the belief value for x should be only 0.6 and 0 for y. The evidence retraction method produces adjusted belief values for R_4:

R_4: if a, b and c then x with -0.1163 and y with -0.7442

In this case, the belief in y should be completely removed from the rule and also belief in x should be slightly increased. Notice that the revised belief values are negative. Though a negative belief value is not allowed in the Dempster-Shafer theory, it is necessary to achieve the intended goal by removing the portion of the effect by regular evidence combination. Negative belief values are not meaningful in any sense for the evidence theory but are used only to attempt to apply the portion of belief in an opposite way.

A more comprehensive evidence combination occurs with 5 rules as follows. The evidence set is densely mixed and spreads over 5 rules forming different sets of conditions. As more conditions are involved in the rule, higher belief values are assigned, and a smaller hypothesis set is assigned. Evidence dependency network for these rules is shown in Figure 3.

R_5:	if a	then x with 0.2 and y with 0.3
R_6:	if a and b	then x with 0.3 and y with 0.5
R_7:	if a and d	then x with 0.5 and y with 0.2
R_8:	if a, b, and c	then x with 0.6
R_9:	if a, b, and d	then y with 0.7

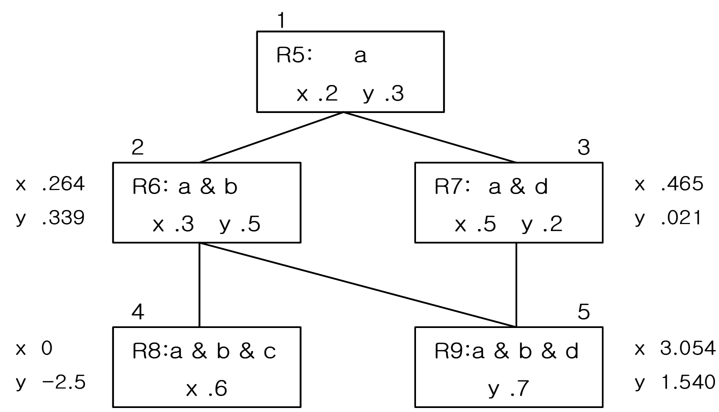

Figure 3. Evidence Dependency Network

Each node corresponds to a rule. Numbers on the top left corner of the box are node numbers. Node 1 corresponds to rule R_5. Two nodes are connected by a link if they have a dependency relation between them. The links are drawn only if two nodes have an immediate subset-superset relation. A link between node 1 and node 5 is not valid because they are not in an immediate subset-superset relation. Belief adjustment can be done on the link between node 1 and node 2 as we have seen above. Adjusted belief values are shown on the side of the box. Note that belief value adjustment should be done in the order from top to bottom. Without adjustment on node 2, no further adjustment is possible below the node. Negative belief values are obtained after adjusting belief values on node 4 based on adjusted values on node 2. Belief adjustment for node 5 needs the two adjustments for node 2 and 3 to be done. Since these two nodes are linked to node 5, the belief adjustment requires regular evidence combination first to get revised belief values. With these revised belief values, the evidence retraction method produces adjusted belief values for node 5. Note also that the adjusted belief values are more than 1.0.

5. Conclusion

The evidence retraction algorithm is presented to bridge the gap between human expert and the system for uncertainty handling. In this paper, the Dempster-Shafer theory is chosen for evidence combination. In the real world, a non-monotonic way of reasoning is common. The evidence retraction algorithm copes with this type of reasoning. Adjusting belief values sometimes leads to negative belief values to reverse the effect of unwanted evidence combination. When regular evidence combination is applied, negative belief values are used to influence belief values in an opposite way.

References

[1] L.S. Dutt and M. Kurian, "Handling of Uncertainty – A Survey," Int. Journal of Scientific and Research Publications, Vol. 3, Issue 1, pp. 1-3, 2013.

[2] K. Sentz and S. Ferson, "Combination of Evidence in Dempster-Shafer Theory," Sandia National Lab., Apr. 2002.

[3] J. Stone, "Bayes's rule: A Tutorial Introduction to Bayesian Analysis," Sebtel Press, England, 2013.

[4] W. Bolstad, Introduction to Bayesian Statistics, John-Wiley, 2007.

[5] R. Yager, "Fuzzy Relations between Dempster-Shafer Belief Structures," Knowledge-based System, Vol 105, pp. 60-67, 2016.
DOI: https://doi.org/10.1016/j.knosys.2016.04.027

[6] J. Merigo, A. Gil-Lafuente and R. Yager, "An Overview of Fuzzy Research with Bibliometric Indicators," Applied Soft Computing, 27, pp. 420-433, 2015.
DOI: https://doi.org/10.1015/j.asoc.2014.10.035

[7] L. Giordano and F. Toni, "Knowledge Representation and Non-monotonic Reasoning," A 25-year Perspective on Logic Programming, pp. 87-111, 2010.

[8] K. Zhou, A.Martin, and Q. Pan, "Evidence Combination for a Large Number of Sources," 20th Int. Conf. on Information Fusion, Jul. 2017, China

[9] G. Lee, "An efficient Dempster-Shafer Evidence Combination Scheme for Uncertainty Handling," The Korea Information Processing Society, Vol. 5, No. 2, pp. 908-914, 2002.

[10] Q. Chen, A. Whitbrook, U. Aickelin, and C. Roadknight, "Data Classification Using the Dempster-Shafer Method," Journal of Experimental & Theoretical Artificial Intelligence, Vol. 26, Issue 4, pp. 493-517, 2014.
DOI: https://doi.org/10.1080/0952813X.2014.886301

[11] Q. Chen and U. Aickelin, "Anomaly Detection Using the Dempster-Shafer Method," Journal of Experimental & Theoretical Artificial Intelligence, Vol. 26, Issue 4, pp. 493-517, 2014.